Evaluating interactive systems from an emotional perspective*

Yenny Alexandra Méndez Alegría,[1] César Alberto Collazos Ordóñez[2]
Universidad del Cauca, Popayán, Colombia.
Antoni Granollers Saltiveri[3]
Universidad de Lleida, Lleida, España.
J. Alfredo Sánchez Huitr[4]
Universidad de las Américas Puebla, Cholula, Puebla, México.

Recibido: febrero 19 de 2014 - Revisado: abril 23 de 2014 - Aceptado: mayo 14 de 2014

Referencia para citar este artículo: Méndez, Y. A., Collazos, C. A., Granollers, T., & Sánchez, J. A. (2014). Evaluating interactive systems from an emotional perspective. *Revista Científica Guillermo de Ockham, 12*(1), 43-49

Abstract

In this paper, we introduce a model that aims to provide guidelines that will strengthen the evaluation of interactive systems by assisting in the identification and analysis of emotions evoked by users during system usage. We briefly discuss related projects that have included emotions in the evaluation of interactive systems. The research presented here is a preliminary work towards the inclusion of emotions during the evaluation of interactive systems. Our model is presented including details of each of its phases. We discussed preliminary results of the application of the model to evaluating a Virtual Learning Environment. Our approach comprises four major phases: Selection of relevant emotions; analysis of relationships between emotions and interactive systems; selection of detection mechanisms; and application of evaluation methods.

Keywords: Emotions, evaluation, user experience, usability.

La evaluación de los sistemas interactivos desde una perspectiva emocional

Resumen

En este trabajo se presenta un modelo que tiene por objeto proporcionar directrices que fortalezcan la evaluación de sistemas interactivos, ayudando así en la identificación y análisis de las emociones evocadas por los usuarios durante el uso del sistema. Se discuten brevemente los proyectos relacionados que han incluido las emociones en la evaluación de sistemas interactivos. La investigación que aquí se presenta es un trabajo previo a la inclusión de las emociones

* This paper is the result of research "Framework for semi-invasive capture and analysis of emotions in the evaluation process of interactive systems with users", supported by the research groups IDIS, GRIHO & ICT. A part of this paper was published in Proceedings of the 2013 Chilean Conference on Human-Computer Interaction.

1. IDIS Research Group. Correspondence concerning this article should be addressed to Calle 26B Norte #6B-40. Barrio Palacé (Diagonal a la Iglesia), Popayán (Cauca). E-mail: ymendal@unicauca.edu.co
2. IDIS Research Group. E-mail ccollazo@unicauca.edu.co
3. GRIHO Research Group. E-mail antoni.granollers@udl.cat
4. ICT Laboratory. E-mail alfredo.sanchez@udlap.mx

durante la evaluación de los sistemas interactivos. Nuestro modelo se presenta con datos de cada una de sus fases y se discuten los resultados preliminares de la aplicación del modelo para la evaluación de un entorno virtual de aprendizaje. El enfoque comprende cuatro grandes fases: selección de las emociones relevantes; análisis de las relaciones entre las emociones y los sistemas interactivos; selección de mecanismos de detección; y la aplicación de métodos de evaluación.

Palabras clave: Emociones, evaluación, experiencia del usuario, usabilidad.

Introduction

Human-Computer Interaction (HCI) is a discipline that (among other aspects) affords with the design of computer interfaces to be used by humans, setting enjoyable to use, attractive and accessible systems (D'Mello & Calvo, 2013). It is important in HCI the user activity that involves a user with a machine; this activity has the physical, cognitive, and affective aspects. The physical level is related with the interaction between computer and human, the cognitive with the different ways that users can understand and interact with the system. Finally, "the affective aspect tries to affect the user in a way that make users continue to use the system by changing attitudes and emotions toward the user" (Saroha, Sharma & Bhatia, 2011).

Emotions are reactions to events related with the concerns, needs or goals of an individual, this includes the components: physiological, affective, behavioral, and cognitive (Brave & Nass, 2008). Only during recent decades it has been recognized that emotions play a key role in all activities that people perform with support from computers (Norman, 2004). This includes searching the web, playing games, exchanging email, buying online, among many others (Peter, Beale, Crane & Axelrod, 2007; Brave & Nass, 2008). Perceiving emotions during system usage can be helpful to determine whether or not the user is satisfied, and that knowledge can be used to adjust and improve subsequent developments (Picard, 2000).

In recent years, an important number of projects (Agarwal & Meyer, 2009; Lera & Garreta-domingo, 2007; Branco, Firth, Encarnação & Bonato, 2005; Petrie, 2010)we explored the use of two techniques: an emotional think aloud protocol and an emotion word prompt list (EWPL have aimed to address the emotional aspects in the evaluation of usability of interactive systems (and to enhance user experience). Objective and subjective aspects of emotions are addressed by various methods used to evaluate interactive systems (Bak, Nguyen, Risgaard & Stage, 2008)

Still, existing methods have significant limitations. Moreover, in spite of the diversity of approaches, a well-defined methodology for incorporating emotions in the process of evaluating usability and user experience remains to be developed.

Classical methods based on questionnaires and video analysis evaluate affect by eliciting user opinions on how they felt prior to and during specific experiments. Users are asked to recall how they feel at each moment as they perform a collection of pre-assigned tasks. More recent approaches take advantage of various sensors placed on the human body (Branco et al., 2005; Tsai, Chen & Lo, 2009).

Thus, in each of these cases emotion identification is based on *a posteriori* user reports, observation by the evaluators, or invasive methods. Even though the detection of affective states with these methods can be precise, emotions experienced by users may have been produced or altered by the very measurement instruments or conditioned by the evaluation context (Picard & Daily, 2005).

Another problem with existing methods is related with the immediate and automatic nature of emotions (Agarwal & Meyer, 2009). Given that only the conscious affective state can be captured by instruments that require relatively extended periods of time, actual emotions experienced by users could be distorted.

Considering that emotions are essential for understanding any activity performed by humans, it becomes crucial to contribute to evaluation processes by enhancing the capture and analysis of emotions experienced by users *while* they interact with systems under evaluation. Emotions must be captured unobtrusively, so the users' performance or their perception of the software being evaluated are not affected.

Related work

– Currently, emotional aspects begin to be considered as an important aspect in the design and evaluation of interactive systems. Agarwal and Meyer (2009)

have proposed a methodology integrating verbal and non-verbal emotion scales so as to consider traditional metrics and information about emotional reactions. This is intended to produce inexpensive, rapid and easy to apply means for measuring emotions and to incorporate them into traditional user studies.

– The extent to which FaceReader, a software tool for analyzing facial expressions, can help in measuring emotions is examined by Zaman y Shrimpton-Smith (2006). They complement the results from using this software with the application of traditional questionnaires. Similarly Cowan, Vigentini y Jack (2009) have observed the relationship between anxiety feelings and how a wiki's usability is evaluated. Yet another study (Champney & Stanney, 2007) reports a method and a process related to emotional profiles and shows how emotions can help usability professionals understand emotional reactions with respect to human-system interactions by identifying factors that impact user experience.

– A set of guidelines has been developed by Lera and Garreta-Domingo (Lera & Garreta-domingo, 2007) for evaluating efficiently and at a low cost the affective states of users by considering their reactions during the process of evaluating interfaces. This is intended to complement analysis based on objective, quantitative data obtained from usability testing and on subjective responses obtained from post-test questionnaires.

– Petrie and Harrison (2009) explored the use of two techniques to study user experiences with websites and other interactive technologies: thinking aloud and a list termed Emotion Words Priming List (EWPL). The former technique was adapted using EWPL to detect affective reactions towards interactive technologies, rather than the typical usability problems. During the application of this method, subjects are requested to express their feelings while they perform a series of tasks (Petrie, 2010)we explored the use of two techniques: an emotional think aloud protocol and an emotion word prompt list (EWPL.

– Cognitive walkthroughs and Metaphors of Human Thinking (MOT) are two traditional methods that have been studied by (Frøkjær & Hornbæk, 2008) in order to compare traditional usability methods with others that consider some psychological aspects in the user. They report MOT tends to detect more issues and has a broader coverage of usability problems.

– In addition to work aimed to modify existing usability methods for considering user emotions during system evaluation, other researchers have focused efforts to develop specific methods for identifying user emotions during system or product use. Some of these methods are briefly discussed next.

– PREMO is a self-report instrument that measures non-verbally both pleasant and unpleasant emotions. This includes 14 animated characters, which portray distinct emotions through dynamic facial, corporal and vocal expressions (Desmet, 2004).

– The Positive and Negative Affect Scale (PANAS) is a psychometric scale designed for measuring the independence of negative and positive affect (Watson, Clark & Tellegen, 1988)positive and negative affect have consistently emerged as two dominant and relatively independent dimensions. A number of mood scales have been created to measure these factors; however, many existing measures are inadequate, showing low reliability or poor convergent or discriminant validity. To fill the need for reliable and valid Positive Affect and Negative Affect scales that are also brief and easy to administer, we developed two 10-item mood scales that comprise the Positive and Negative Affect Schedule (PANAS. PANAS was developed mainly for clinical situations but is also utilized in evaluation studies in which the users' mood could be affected. The scale can be applied both to prototypes and to commercially available products (Thompson, 2007).

– Emocards (Desmet, Overbeeke & Tax, 2001) is an instrument that consists of 16 animated faces that represent eight different emotional responses. This instrument assumes that each idea can be associated with a specific recognizable facial expression.

– These works represent strides towards the integration of emotional aspects into the evaluation of interactive systems. It is important to note, however, that subjectivity continues to be an issue when analyzing their results, as they rely on instruments such as questionnaires or direct observation by evaluators.

Our model

In this section, we present our model, aimed to include emotions in the process of evaluating interactive systems. The model is based on Emo-Eval project (Méndez, Collazos, Granollers & Sánchez, 2013) and includes four

major phases, which will be detailed in next subsections: (1) *selection of relevant emotions; (2) analysis of relationships between emotions and interactive systems; (3) selection of detection mechanisms; and (4) application of evaluation methods.*

Selection of relevant emotions

In this phase, we consider activities related with this question: *What are the most important emotions to consider during the process of evaluating specific interactive systems?* This aspect allows for the identification of the most relevant emotions we required in order to delimit the research scope. Although many works have already identified and classified many emotions, it is a good idea to choose emotions that arise in the specific context of interactive systems evaluation. This selection strongly depends on the specific interactive system under evaluation. For example, if we are working with a bank web portal, maybe it is not necessary to include emotions such as desire or disgust, among others. But, if we are working with a game, probably it will be necessary to consider emotions such as frustration, disgust, and surprise. Some factors that could help determining the most relevant emotions in each case include the purpose of the evaluation, the available or necessary resources, and the estimated time for their detection and interpretation.

Analysis of relationships between emotions and interactive systems

In this second phase, we include activities related with this question: *How emotions should be analyzed during the evaluation process?* This implies the consideration of these aspects:

(1) Identifying the relationship between the actions performed by the users when interacting with the system and the emotions selected in the previous phase; and

(2) Establishing metrics to provide an objective assessment of the emotions selected.

(3) Interpreting and analyzing the information that has been previously identified.

Selection of detection mechanisms

During this phase we consider elements related with: *What are appropriate mechanisms for detecting and identifying relevant emotions?* Once emotions have been selected, it becomes critical to opt for feasible mechanisms for dealing with emotions in terms of effectiveness and efficiency. Three activities are related with that question:

(1) Establishing subjective mechanisms that are based on user responses after system usage (e.g., questionnaires, interviews, etc.) to identify emotions;

(2) Establishing mechanisms related with physiological manifestations of emotions. Progress in this area includes various alternatives to measure heart rate, breathing rate, galvanic skin, among others; and.

(3) Establishing mechanisms related with the identification of emotions based on gesture and voice recognition.

The main purpose of this phase is to adapt the detection and measurement mechanisms depending of the selected emotions. For example, emotions in PREMO have been divided into seven positive and seven negative emotions. Positive emotions are: inspiration, desire, satisfaction, pleasant surprise, fascination, amusement, and admiration. And, negative emotions are: disgust, indignation, contempt, disappointment, dissatisfaction, boredom, and unpleasant surprise.

Some emotions in PREMO are not important for all interactive systems. For example, as noted earlier, for a bank web portal application, an instrument based on PREMO probably should not include emotions such as desire, fascination and amusement, as they may not be relevant during the evaluation. However, probably for a game evaluation, all the emotions established in PREMO should be considered.

Application of evaluation methods

Activities to be considered during the evaluation phase include the following:

(1) Planning activities in order to "neutralize emotions". In other words, it is necessary give to the participants of the evaluation, some activities that allow them to "relax" before beginning the user study. Subjects in a study might experience emotions that are not necessarily related with the use of the interactive system under evaluation. Hence the importance of defining alternatives for activities to "neutralize" previously existing emotions.

(2) Defining an appropriate scenario for conducting the study. This is important as emotions evoked by par-

ticipants could be affected by factors that are external to the sole use of the system being evaluated.

(3) Configuring mechanisms for emotion detection that were selected in the previous phase.

(4) Selecting appropriate personnel for coordinating and supporting the evaluation process.

(5) Selecting evaluation methods, which typically will depend on the stage of the development of the system to be evaluated (non-functional or functional prototype, specific functionality or completed system).

The model we have defined has been applied to the incorporation of emotions in the context of evaluating Virtual Learning Environments (VLE). Preliminary results are presented from applying three of the Emo+Eval phases: *Selection of relevant emotions, selection of detection mechanisms, and application of evaluation methods.*

Selection of relevant emotions for VLEs

Considering that relevant emotions vary according to the specific system to be evaluated, we have studied the kinds of emotions evoked by students in VLEs. In particular, we are interested in emotions that may motivate students to utilizing the VLE and undertaking the tasks assigned to them through the system.

From our review of the literature (Reeve, 1996; 2010; Damasio, 2003), we have found that the following emotions are relevant for the evaluation of VLEs:

– *Inspiration.* If a VLE offers alternatives for the students to become inspired, they may get motivated to continue to use the system and perform their assigned tasks.

– *Amusement.* VLEs should be designed in a sense that encourages users to find entertaining and pleasing ways to learn. This will vary according to factors such as age and cultural background.

– *Satisfaction.* Satisfied students will be highly motivated. VLEs should be developed in ways that instructors can design materials that promote student satisfaction.

– *Disillusion.* Any form of disillusion with the interface or the contents in the VLE will affect motivation negatively and discourage students from taking full advantage of the system.

– *Anguish.* A closely related to frustration emotion, becoming a key emotion, that should be avoided in VLEs or in any form of learning.

– *Anger*. This emotion may be derived from poor design features, and is also directly related to frustration.

– *Indignation.* VLE designers should be aware that a patronizing tone or a text that implies prejudices may prompt this emotion and impact motivation and learning.

Selection of detection mechanisms for VLEs

We have selected subjective mechanisms for emotion detection, which include questionnaires and interviews with eight students, recruited from one of the participating universities where VLEs are under evaluation.

We also have been working closely with a group doing research in the area of gesture and voice analysis (Cortés-Silva, 2013). Our plan is to take advantage of progress in this area in order to strengthen the results obtained through subjective instruments.

Application of evaluation methods for VLEs

The following are activities we have undertaken as part of this phase:

Scenario configuration. We used an eye tracker device together with VLE. Online questionnaires were prepared so they could be applied prior to and after tasks were performed.

Emotion neutralization. Our participants started viewing a set of sequences of photographs suggested by a psychologist (seascapes, forests and blossoming gardens). Relaxing music has also been suggested as an alternative or a complementary stimulus. *Task execution*. Students were asked to start (or respond to) a discussion using a forum in the VLE.

Subjective data collection. An online PREMO-based questionnaire was applied to our participants.

Analysis of results. This activity is currently ongoing. We are identifying the emotions evoked by students and will analyze their relationships with the VLE. We plan to undertake two initial activities:

– *Evaluation by experts in emotions.* Emotions will be identified by presenting images to a group of experts who may be behavior professionals or actors. Images will include faces taken from videos recorded during task execution. Emotions will be determined (and associated to specific instants) from interpretation and consensus reached by the experts.

– *Report results*. In addition to the emotions determined by the group of experts in emotions, we will report their intensity and their relationships with each of the tasks executed by participants.

Conclusions and further work

We currently are working on refining the tasks comprised by our model. We aim to provide alternatives that make it feasible for evaluators to accomplish each or the activities that are included in Emo+Eval's phases. In addition to completing the VLE case study, the model will be applied to other application domains. We are planning to study its effectiveness in the context of innovation processes supported by multi-tactile interactive surfaces. We also are considering the use of physiological sensors to corroborate the experts' judgments on evoked emotions. This will allow us to improve the activities that make up each of the phases of Emo+Eval. We consider this model as a starting point for strengthening the process of evaluating interactive systems by providing general guidelines for considering emotions as part of that process.

Acknowledgments

The authors gratefully acknowledge support for this project from Colombia's Administrative Department of Science, Technology and Innovation. We also would like to acknowledge the enthusiastic collaboration from the GRIHO research group at the University of Lleida, the ICT Lab of University of the Américas Puebla, and the Program for Academic Mobility between Andalusian and Latin American AUIP Universities.

References

Agarwal, A., & Meyer, A. (2009). Beyond usability: evaluating emotional response as an integral part of the user experience. In *CHI 09 Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 2919-2930). ACM. doi:10.1145/1520340.1520420

Bak, J., Nguyen, K., Risgaard, P., & Stage, J. (2008). Obstacles to usability evaluation in practice: A survey of software development organizations. *Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges.* Retrieved from http://doi.acm.org/10.1145/1463160.1463164

Branco, P., Firth, P., Encarnação, L. M., & Bonato, P. (2005). Faces of emotion in human-computer interaction. *Work*, 1236-1239. doi:10.1145/1056808.1056885

Brave, S., & Nass, C. (2008). Emotion in Human-Computer Interaction. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications.* (Second., Vol. 29, p. 1386). Standford University: Lawrence Erlbaum Associates. Retrieved from http://courses.washington.edu/imt540a/readings/Blomberg-08.pdf

Champney, R. K., & Stanney, K. M. (2007). Using Emotions in Usability. *Proceedings of the Human Factors and Ergonomics Society*, *2*, 1044-1049. Retrieved from http://www.ingentaconnect.com/content/hfes/hfproc/2007/00000051/00000017/art00012

Cortés-Silva, X. (2013). *Vikara: Plataforma de desarrollo y gestión de aplicaciones para detección de emociones*. Department of computing, electronics and mechatronics. Tesis no publicada. Universidad de las Américas, Puebla.

Cowan, B. R., Vigentini, L. & Jack, M. A. (2009). Exploring the effects of experience on wiki anxiety and wiki usability: an online study. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers Celebrating People and Technology*, 175-183. Retrieved from http://portal.acm.org/citation.cfm?id=1671032

D'mello, S. & Calvo, R. A. (2013). Beyond the Basic Emotions: What Should Affective Computing Compute? In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2287-2294). New York, NY, USA: ACM. doi:10.1145/2468356.2468751

Damasio, A. R. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. *Orlando Harcourt* (Vol. 1st, p. 355). Harcourt. Retrieved from http://www.loc.gov/catdir/description/har021/2002011347.html

Desmet, P. (2004). Measuring emotion: development and application of an instrument to measure emotional responses to products. In M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology* (Vol. 41, pp. 111-123). Kluwer Academic Publishers. doi:10.1007/1-4020-2967-5

Desmet, P., Overbeeke, K., & Tax, S. (2001). Designing Products with Added Emotional Value: Development and Appllcation of an Approach for Research Through Design. *The Design Journal*, *4*(1), 32-47. doi:10.2752/146069201789378496

Frøkjær, E., & Hornbæk, K. (2008). Metaphors of human thinking for usability inspection and design. *ACM Tran-*

sactions on Computer-Human Interaction, *14*(4), 1-33. doi:10.1145/1314683.1314688

Lera, E. De, & Garreta-Domingo, M. (2007). Ten Emotion Heuristics : Guidelines for assessing the user's affective dimension easily and cost-effectively. *Group*, *2*(September), 7-10. Retrieved from http://portal.acm.org/citation.cfm?id=1531450

Méndez, Y. A., Collazos, C. A., Granollers, T., & Sánchez, J. A. (2013). Emo+Eval: Including Emotions in the Process of Evaluating Interactive Systems. In *Proceedings of the 2013 Chilean Conference on Human - Computer Interaction* (pp. 48-49). New York, NY, USA: ACM. doi:10.1145/2535597.2535612

Norman, D. A. (2004). *Emotional Design: Why We Love (or Hate) Everyday Things*. *The Journal of American Culture* (Vol. 2006, p. 272). Basic Books. doi:10.1111/j.1537-4726.2004.133_10.x

Peter, C., Beale, R., Crane, E., & Axelrod, L. (2007). Emotion in HCI. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 2* (pp. 211-212). Swinton, UK, UK: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=1531407.1531471

Petrie, H. (2010). Measuring User Experience of websites : Think aloud protocols and an emotion word prompt list. *Methods*, 3673-3678. doi:10.1145/1753846.1754037

Petrie, H. & Harrison, C. (2009). *Measuring users' emotional reactions to websites*. Proceedings of the 27th international conference extended abstracts on Human factors in computing systems (pp. 3847-3852). Boston, MA, USA: ACM. doi: http://doi.acm.org/10.1145/1520340.1520582

Picard, R. (2000). Affective Perception. *Communications of the ACM,*, *43*, 50-51. Retrieved from http://affect.media.mit.edu/pdfs/00.picard-cacm.pdf

Picard, R., & Daily, S. (2005). Evaluating affective interactions: Alternatives to asking what users feel. *CHI Workshop on Evaluating Affective …*. Retrieved from http://mit.sustech.edu/NR/rdonlyres/Media-Arts-and-Sciences/MAS-630Spring-2008/Readings/picard_chi2005.pdf

Reeve, J. (1996). *Motivacion y Emoción* (McGraw-Hil.). University of Rochester, River Campus.

Reeve, J. (2010). *Motivacion y Emoción* (5ª Edi.). México: McGraw Hill.

Saroha, K., Sharma, S., & Bhatia, G. (2011). Human Computer Interaction: An intellectual approach. *IJCSMS International Journal of Computer Science and Management Studies*, *11*(02), 147-154. Retrieved from http://www.ijcsms.com/journals/Volume 11, Issue 02, Aug 2011_Paper25.pdf

Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227-242. doi:10.1177/0022022106297301

Tsai, T. C., Chen, J. J., & Lo, W. C. (2009). Design and Implementation of Mobile Personal Emotion Monitoring System. In *Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware* (pp. 430-435). Washington, DC, USA: IEEE Computer Society. doi:10.1109/MDM.2009.78

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development of brief measures of positive and negative affect: The PANAS scale. *The PANAS Scales Journal of Personality and Social Psychology*, *47*, 1063-1070.

Zaman, B., & Shrimpton-Smith, T. (2006). The Face Reader: Measuring instant fun of use. In A. Mørch, K. Morgan, T. Bratteteig, G. Ghosh, & D. Svanaes (Eds.), *NordiCHI* (Vol. 189, pp. 457-460). ACM Press. doi:10.1145/1182475.1182536